

A Case Study of Return on Investment for Multi-sites Test Handler in The Semiconductor Industry Through Theory of Industry 4.0 ROI Relativity

<https://doi.org/10.3991/ijes.v7i3.11057>

Khoo Voon Ching

University Technology of Malaysia, Kuala Lumpur, Malaysia
khoo.vc@gmail.com

Abstract—The Fourth Industrial Revolution started in the year 2011 with the aim to improve productivity similar with that of the previous three industry revolutions that occurred 200 years ago. The revolutions were promoted and implemented to improve the efficiency and speed of productivity. However, Industry 4.0 is likely a supplementation of the existing approach, with the purpose to centralize the processes and factories together to be controlled by a central console. The goal is to procure productivity and manufacturing data to enable data analysis so that the industry's performance affected by big data variables, namely, velocity, veracity, variety, value, and visibility are monitored and rectified for continued productivity. This research initiates the development of a theory of Industry 4.0 ROI Relativity from the economic theory of firms, incorporating the pick-and-place test equipment and Industry 4.0 variables. A case study through the experimental research approach (ERA) was conducted by measuring the effects of velocity and veracity accuracy on the good-unit per hour (UPH), profit, and return on investment (ROI) of the pick-and-place test equipment and Industry 4.0. Then, the data were analyzed with Pearson correlation coefficient to determine inter-correlations among the velocity and veracity accuracy with the UPH, profit, and ROI. This research concluded that a significant average and negative correlation exists among velocity, UPH, profit, and ROI. Furthermore, the inter-correlation analysis results show a significant average and positive correlation among veracity accuracy percentage, UPH, profit, and ROI.

Keywords—Cost of Test; Testing economic model; Multi-site testing; Industry 4.0; Return on Investment Model; Big Data analysis.

1 Introduction

Industrial revolutions have begun approximately 200 years ago in Great Britain, with the goal to improve productivity. Before the revolutions, manufacturing processes were 100% manually produced by human labor. The first industrial revolution was implemented by replacing human labor with mechanical looms driven by steam in the year 1780 [8] and it was labeled as the “First Industrial Revolution” [1]. The first

industrial revolution quickly increased productivity because the speed of manufacturing increased through the mechanical looms driven by steam and water machine when compared with human labor. Hence, the Industrial output increased by an annual average of 2.46% between the years 1780 and 1880 [6], as shown in Figure 1 below. When the Second Industrial Revolution began nearly 100 years later, adopting the concept of division of labor, different processes were lined up in one manufacturing linear-line from station to station until it produced the final goods. The Second Industrial Revolution had enhanced the division of labor which emerged from the classical theory, the latter with the aim of saving time that was often lost when work passed from one person to another [25]. In the technology era, division of labor can be identified as the division of processes or machines. This process can be conducted by linking all processes with a conveyor powered by electricity, substantially reducing station-to-station transfer times. With this invention, the industry output growth peaked at an average of 7.24% per year. Then the Industrial Revolution 3.0 that followed produced a significant growth in industry output at an annual average of 51.83%. This growth was attributed to the automated manufacturing line employing technology digitization in the industry through the implementation of computer approach and robotic automation.

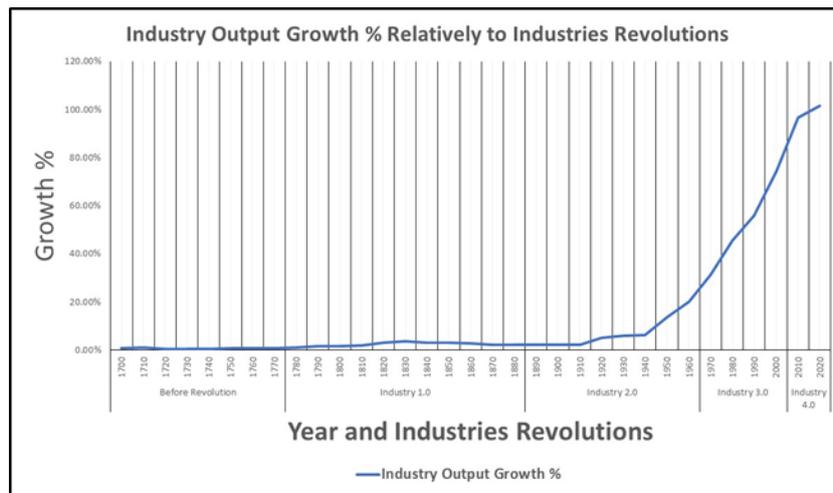


Fig. 1. Industry Output Growth Percentage Relative to Industry Revolutions

The history of industrial revolutions proved that industry output growth has been consistently generated by the implementation of new approaches, as shown in Figure 1 above. However, Industry 4.0 is very different than its predecessors. Three revolutions were invented and implemented to improve efficiency and speed of productivity. Industry 4.0 is likely a supplementation of the existing approach, with its aim to centralize processes and factories together, controlled by a central console to obtain productivity and manufacturing data which are subsequently analyzed to determine required actions that will improve processes and reduce errors caused by inefficiency

in data collection, communication, and data analysis. In this way, productivity can be improved to further enhance profit. The power of 1% is able to improve the production output and ultimately appreciate profit [11]. For example, in the aviation industry, fuel savings of 1% per annum is equal to savings of 30 billion dollars. In short, the Industry 4.0 revolution plays the supporting role of assisting existing technologies and processes to improve their efficiency. However, it cannot work by itself.

The basic concept of Industry 4.0 is to connect the entire manufacturing equipment and processes with the central console unit to enable machine-to-machine communication through the central console unit. In addition, manufacturing data are retrieved from physical machines through feedback devices such as sensors and vision-inspection systems. Then, the data are managed and stored in cloud technologies such as Hadoop, an open source cloud-based distributed data storage system. The data will then undergo analysis via technologies, and decisions will be made through artificial intelligence (AI) and machine learning (ML) capabilities to improve productivity, reduce rejection rate, and increase utilization percentage by reducing unnecessary or unplanned downtime. This approach is called 'cyber-physical system'. The entire concept will only work if big data analysis is enabled. In other words, big data analysis is the brain of Industry 4.0, whereas the Internet of Things, such as server and data processing devices, are the heart of Industry 4.0.

The implementation of big data in Industry 4.0 is determined by five variables (Vs), namely, velocity as the server communication and processing speed, veracity as the trustworthiness of the data, variety as the different types of data reflecting the system's capability to process, value as the data type to provide the most important information for certain processes, and visibility of data that enables end users to easily understand information.

2 Problem Statement

All industry revolutions aimed to improve productivities. Industry 4.0 is no exception but it is different in its purpose, aim and approach to achieve improved productivity. The latter is achievable when the centralized-connectivity control-console revolutionizes the processes of target manufacturing productivity. However, the question remains if this implementation genuinely will benefit the industry or not? Big data are the brain of Industry 4.0, but such volumes exceed the capacity of the current on-line storage and processing system and thus cause uncertainty [14]. Moreover, the five Vs are the determinants of Industry 4.0 to improve productivity and hence increase profit as the ultimate goal. Therefore, investigating the five Vs is important to verify if Industry 4.0 is the solution to manufacturing issues in the future since its introduction in 2011. Most of the styles are intuitive. However, we invite you to read carefully the brief description below.

The five big data Vs are variables that directly contribute to the Industry 4.0 revolution. Many researchers have put all of them side by side. Nonetheless, the authors have categorized them into two categories, namely, primary and secondary variables, suggesting that the latter, namely, value, visibility, and variety, are the contributors to

velocity and veracity; and the former includes the following: (1) Data velocity measures the speed of data creation, streaming, and aggregation [14] and it might affect the manufacturing processing time. The communication times must be close to zero to ensure that it will not slow down the existing manufacturing processing time [11]. This possible scenario is due to the fact that velocity may affect productivity positively and negatively [12]. Figure 2 shows the different types of servers producing different communication speeds. Basically, two types of database servers exist, namely, SQL and NoSQL, which produce different velocities. SQL is a conventional database developed and owned by Microsoft, which adopts the relational model. By contrast, non-relational databases, commonly known as NoSQL database, were developed to improve processing and communication velocity [21]. (2) Variety is a truthful and credible presentation of data, where the data by default can be unreliable and downright false [11]. The tradeoff in big data set is determined by the quality of information and cost. IBM estimates that poor data quality costs US consumers by approximately USD 3.1 trillion per year [23]. This finding is an important limitation of the current big data research and practice because without identifying big data veracity, big data-driven discoveries are questionable. Furthermore, the uncertainty in big data veracity was predicted to rapidly increase each year, which comes from multiple sources such as data inconsistency and incompleteness, ambiguities, latency, deception, and model approximations [23]. Big data have no value unless they can be effectively and properly utilized, depending on recognizing and accounting for its information quality [14].

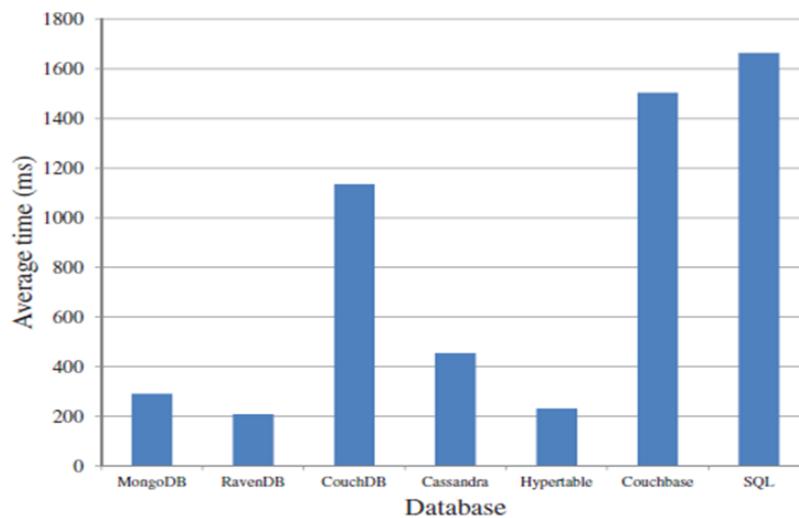


Fig. 2. Database Bucket Instantiating Time (TIDB)

The secondary variables are those that affect the primary variables, namely, (a) data variety, which measures the richness of data representation, such as texts, images, videos, and audios [14]; (b) data value, which measures the details and importance of

the data contributing to the particular cases for the betterment of the business value [11]; and (3) data visibility, or data complexity, which measures the degree of interconnectedness and interdependence in big data structures [14].

With the aim to improve productivity, and when productivity improves with Industry 4.0, the average cost decreases which, in turn, will then contribute to increasing profit margin. By contrast, if the improvement of productivity does not exceed the capital investment expectation, then the cost of production will be negatively affected [28]. Therefore, this research takes the initiative to fill in the gap by developing the ROI cost model to measure the ROI of Industry 4.0 premised on the five Vs. This will provide a new understanding on the potential contributions of Industry 4.0 to productivity improvement which will surely affect the cost and profit in the consolidated manufacturing processes, the results of which will determine if it indeed an industry revolution has materialized. Hence, this research focuses on semiconductor multi-site test equipment to prove the impending revolutionary change.

3 Development of Theory of Industry 4.0 ROI Relativity for the Semiconductor Multi-sites Test Equipment

The model of Industry 4.0 ROI relativity (TI.4ROIR) is the result of the author's previous research on the ROI of multi-site testing in the semiconductor industry. Previous models, such as the cost of test profit margin model (COTPM) and ROI test technology model (ROITM), have adopted firm average cost, profit, and return on investment (ROI) theories. The present research further aims to improve the model by integrating the Industry 4.0 variables, as shown in Figure 3. The secondary variables, namely, variety, value, and visibility, affect veracity, whereas visibility and variety affect velocity. Veracity is a variable that provides information for AI and ML in Industry 4.0. Here, data veracity affects manufacturing and processing time, and AI determines which test parameters are constantly stable and which can be disabled for the next device testing until it becomes a consecutive failure. Then, the process will resume to those particular parameters. In this case, the testing time can be significantly reduced. However, if the decision from the AI is inaccurate, then the reject yield will increase. Furthermore, the control of the test manufacturing environment will prevent rejects caused by the electrostatic discharge (ESD) because semiconductors are sensitive to ESD. Unplanned equipment downtime can be significantly reduced from the prediction of big data equipment lifespan and wear-and-tear control. However, it will only be achievable if veracity is accurately predicted. AI and ML will not function without feedback from the testing variables, namely, manufacturing processing time, equipment downtime, good-unit yield, and environment control. Again, big data veracity must be accurate. As shown in Figure 3, this process must be a two-way communication between Industry 4.0 and the automated manufacturing line to enable AI and ML.

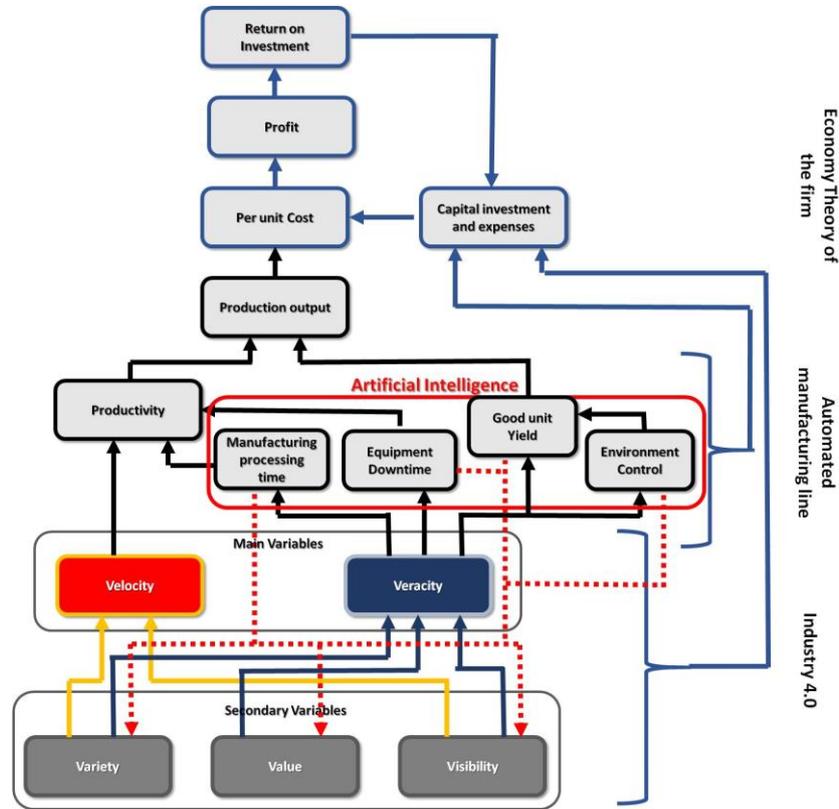


Fig. 3. Model of Industry 4.0 ROI relativity

By contrast, velocity affects productivity, as indicated in the previous section. Communication time must be close to zero to ensure that it will not slow down the existing manufacturing processing time [11]. In other words, big data velocity will be an additional processing time on top of the existing processing time. This method will slow down the total manufacturing processing time than the one that does not implement Industry 4.0.

Good-unit yield and productivity will directly affect production output. According to firm average cost theory, if the production output decreases, then the cost per unit will increase. The latter affects profit margin which is the main variable contributing to ROI. Thus, we have developed the model of TI.4ROIR by integrating the COTPM and ROITM with the five big data Vs.

$$\text{Profit} = (\text{Demand} \times \text{Selling Price}) - \left(\frac{\left(\frac{\text{Dep} + \text{DL} + \text{OH} + \text{CPKG}}{729.6} \right) \times \text{NOTE}}{\frac{\text{Total Good device}}{\text{Total Incoming Device}} \times 100\%} \right) \quad (1)$$

Equation 1 is the cost of test profit margin model developed from the profit theory [16], where the profit is equal to the total revenue minus the total cost. The variables shown in Equation 1 are explained as follows:

- Demand is the total unit per hour (UPH), as shown in Equation
- Selling price is the average selling price of a semiconductor chip
- Dep is the depreciation cost per month of the test equipment
- DL is the direct labor cost per month
- OH is the overhead cost per month
- CPKg is the bad-unit cost
- N is the number of test sites
- t is the test time
- i is the indexing time
- 729.6 is the number of hours per month
- NOTE is the number of test-equipment setups required for a particular production output
- Actual utilization percentage (AU%O) can be calculated by dividing the utilization percentage per output (U%O) by the number of test-equipment setups necessary to produce the required test output

$$\text{Demand} = \text{Unit per hour} = \left(\frac{3600 \times N}{(1-MSE) \times (N-1) \times (t1+i1) + (t1+i1)} \right) \times \left(\frac{TGD}{TID} \times 100\% \right) \quad (2)$$

where

- 3,600 is the number of seconds in an hour
- N is the number of test sites
- MSE is the multi-site efficiency
- i1 is the single-site indexing time
- t1 is the single-site test time
- TGD is the total good devices
- TID is the total incoming devices

Equation 2 needs to further improve the inclusion of the Industry 4.0 variables. With reference to Figure 3, aside from the indexing time and testing time, big data velocity time must also be considered. Furthermore, as shown in Figure 3, testing yield is affected by environment control, AI, error of the testing yield, and environment control, which can be eliminated if the prediction of veracity is accurate. Finally, the test time will be decreased by the AI control to disable some of the test parameters when they are discovered as stable or “reduce pin count test,” to improve testing throughput [27]. Therefore, Equation 3 was developed to incorporate the two Industry 4.0 variables, namely, velocity and veracity, and replace the MSE with the AICMSE:

$$\text{Good unit per hour 1.4} = \left(\frac{3600 \times N}{((1-AICMSE) \times (N-1) \times (t1+i1) + (t1+i1) + I.4Vt)} \right) \times AIY\% \quad (3)$$

where

- I.4Vt is the velocity time, as shown in Equation 4
- AIY% is the AI yield, as shown in Equation 5
- AICMSE is the AI-corrected multi-site efficiency, as shown in Equation 6

$$I.4Vt = \text{(Time for Instantiating Database bucket } X \text{ Number of test parameter)} X 2 \quad (4)$$

As shown in Equation 4, I.4Vt is equal to the database bucket instantiating time (TIDB) multiplied by the number of test parameter (NOTP), where the NOTP is the number of parameters verified by the semiconductor tester on a semiconductor chip. Furthermore, the sum of the velocity time must be multiplied by two because, as shown in Figure 4, two-way communication exists. Other data storing methods are also available, such as flat files, ISAM, head files, and B+ trees. This study focuses on the hash bucket method.

$$AIY\% = ((100\% - (AY\% - ECR\%)) X VRA\%) + (AY\% - ECR\%) \quad (5)$$

As shown in Figure 3, the testing yield is affected by environment control (ECR%) and veracity accuracy (ARA%). Therefore, as shown in Equation 5, AIY% is mean AI-corrected yield percentage, which is equal to 100% (total incoming devices) minus the actual test yield percentage (AY%) deducted in the environment control error %, which causes rejection owing to ESD. Then, the difference is multiplied by the veracity accuracy percentage (VRA%). The first portion of the equation calculates how much yield improvement will be contributed via the AI, which is affected by the data veracity accuracy. Furthermore, the value obtained from the first portion will then be added into the second portion, which is the actual test yield percentage (AY%) deducted in the rejection caused by the environment control (ECR%).

$$MSE = 1 - \left[\frac{(tms+ims)-(t1+i1)}{(N-1)x(t1+i1)} \right] \times 100 \quad (6)$$

As shown in Equation 6, the multi-site efficiency equation cited from [16] and developed from [15] is the MSE equation, where tms and ims are the test time and indexing time for multi-sites, respectively; t1 and i1 are the test time and indexing time for single site testing, respectively; and N is the number of test sites. To calculate the MSE, which included the AI-corrected multi-site test time and depended on the VRA%, tms needs to be further computed, as shown in Equation 7.

$$AICtms = [TTP - (TTP x VRA\%)] x \left(\frac{ATT}{TTP} \right) \quad (7)$$

As shown in Equation 7, AICtms is equal to the total test parameter (TTP) minus TTP and multiplied by the veracity accuracy percentage (VRA%). If VRA% is equal to zero, then the entire TTP will be tested. If the VRA% increases, then the TTP will decrease, which is also known as reduce pin count test. The higher the VRA%, the lesser the test parameter will be tested. It will then be multiplied by the actual test time (ATT) deducted by TTP to obtain the average test time per test parameter.

The computation of AICMSE is shown in Equation 8, which replaces tms with AICtms.

$$AIMSE = 1 - \left[\frac{(AICtms+ims)-(t1+i1)}{(N-1)x(t1+i1)} \right] \quad (8)$$

Another variable shown in Figure 3 is equipment utilization. Through the AI control, unplanned downtime shall be reduced but still depends on the data veracity accuracy (VRA%). Therefore, as shown in Equation 9, I.4AU%O is derived.

$$I.4AU\%O = AU\%O - [AU\%O - (AU\%O \times VRA\%)] \quad (9)$$

As shown in Equation 9, the equipment downtime computation is derived by AU%O minus AU%O multiplied by VRA%. If VRA% is equal to 100%, then equipment downtime will be equal to zero. However, if VRA% decreases, then equipment downtime will increase.

The final profit equation, which incorporated the Industry 4.0 aspects, can be derived as shown in Equation 10: (a) the demand replaced by the good UPH 1.4 as shown in Equation 3, (b) the testing yield replaced by the AIY% as shown in Equation 5, (c) AU%O replaced by I.4AU%O as shown in Equation 9, and (d) Industry 4.0 cost (I.4\$) added into the part of the total cost on top of Dep, DL, OH, and CPKG costs.

$$Profit = (Good\ unit\ per\ hour\ 1.4 \times Selling\ Price) - \left(\frac{\left(\frac{Dep+DL+OH+CPKG+I.4\$}{729.6} \right) \times NOTE}{I.4AU\%O} \right) \times AIY\% \quad (10)$$

Furthermore, to develop the ROI model, the relation of ROI with cost and profit must be identified, as shown in Equation 11. The computation of ROI was conducted by dividing the net income with the investment [26], where the net income is equal to revenue (R) minus the cost (C), which is the profit. The profit for this study was calculated in Equation 10. The investment of the semiconductor multi-site testing included depreciation cost (Dep), direct labor cost (DL), overhead cost (OH), bad-unit cost (CPKG), and Industry 4.0 cost. Therefore, the equation that included the Industry 4.0 aspect is derived, as shown in Equation 12.

$$ROI = \frac{Revenue(R) - Cost(C)}{Investment(I)} \quad (11)$$

$$TI.4ROI = \frac{(Good\ unit\ per\ hour\ 1.4 \times Selling\ Price) - \left(\frac{\left(\frac{Dep+DL+OH+CPKG+I.4\$}{729.6} \right) \times NOTE}{I.4AU\%O} \right) \times AIY\%}{\left(\frac{Dep+DL+OH+CPKG+I.4\$}{729.6} \right)} \quad (12)$$

As indicated in Equation 12, the investment and cost must be divided by 729.6 to obtain per hour cost because the demand/output of the equation is set for one hour. In the semiconductor industry, 729.6 is the number of hours set for a month, which is calculated by using 52.112 weeks per year divided by 12 to obtain the number of weeks in a month. Then, the quotient is multiplied by seven days to obtain the average number of days per month and then multiplied by 24 hours to obtain the number of

hours per month. For a standard calculation, one hour is equal to 3,600 seconds [9]. Furthermore, depreciation cost (Dep) is the cost of the test equipment, including the test handler and tester, divided into five (on the basis of the life span of five years) to obtain the Dep per year and is further divided into 12 to obtain the cost per month. All other costs, namely, direct labor cost (DL), overhead cost (OH), and bad-unit cost (CPKG and Industry 4.0 cost [I.4\$]), are computed as one-month costing. When divided by 729.6, it will become a per hour cost. The author of this research names Equation 12 as theory of Industry 4.0 ROI Relativity (TI.4ROI_R).

4 Research Methodology

4.1 Pick-and-place test equipment variables

An experimental research was conducted on the multi-site pick-and-place test equipment (PNPTE), where the variables for the PNPTE are tabulated in Table 1.

Table 1. Research variable setting

| PNPTE variables | |
|--------------------------------------|-----------|
| Number of test-sites (<i>N</i>) | 32 |
| Utilization percentage | 95% |
| Environment control error percentage | 5% |
| Depreciation cost per month (RM) | RM 76,992 |
| Direct labor cost per month (RM) | RM 7,843 |
| Overhead cost per month (RM) | RM 56,937 |
| Industry 4.0 cost per month (RM) | RM 1,567 |

- Depreciation cost (Dep) refers to the test-equipment cost spanning over five years, which comprises the economic life of the product from the purchase price until it reaches a zero value. Industrial measurement standard stipulates that the life span of equipment is five years (at 20% depreciation cost per year).
- Direct labor (DL) cost is the monthly salary of employees, including operators and technicians, who immediately make contributions to the production output.
- Overhead (OH) cost is the cost incurred during manufacturing, aside from equipment depreciation and direct exertion costs. The components of overhead cost include management cost, facility cost, floor-space cost, maintenance cost, and cost-of-test accessory.
- Industry 4.0 cost is the monthly license cost for every equipment, where the standard charge of the license is USD 4,700 per year per test equipment, which is equal to approximately RM 18,800 per year or RM 1,567 per month.

4.2 Data collection

A total of 12 conditions, six for velocity and six for veracity, were simulated and are tabulated in Table 2.

Table 2. Case study conditions

| Industry 4.0 Variables | Conditions—TIDB |
|------------------------|--|
| Velocity | RavenDB (0.2 second) |
| | MongoDB (0.3 second) |
| | Cassandra (0.45 second) |
| | CouchDB (1.15 second) |
| | SQL (1.6 second) |
| | <i>Conditions—veracity accuracy percentage</i> |
| Veracity | 15% accuracy |
| | 35% accuracy |
| | 55% accuracy |
| | 75% accuracy |
| | 95% accuracy |

To obtain sufficient data for the cost of test, profit margin, and ROI, 30 sets of production lots for X32 test-site configuration were collected as the case study data size (the collected data type is discussed in the following section). Each set of data contained 100 trial runs. Thus, the 30 sets of data contain 3,000 trial run data to obtain test yield, indexing time, and test time. As indicated in Table 1, this research fixed the utilization percentage as 95% and environment control error percentage as 5% to enable a fair comparison of all the conditions as tabulated in the table. The pick-and-place test equipment data, which included test yield, indexing time, and test time, were then calculated with UPH, profit, and ROI via Equations 3, 10, and 12, respectively. The conditions included server TIDB and veracity accuracy percentage as tabulated in Table 2. As a result, the UPH, profit, and ROI affected by Industry 4.0 were obtained.

For the pick-and-place handler, the data collection only considered pure indexing time and rejected any indexing time slowed down by external factors, such as carrier transfer process, loading and unloading processes, equipment jamming, and delay caused by slowing down of pick-arm 1. Production data were accepted only if no external factors are found, such as handler downtime and tester downtime. This study only focuses on the elements within the circle shown in Figure 4.

The indexing time was considered valid only if no waiting time exists between the exchange times for the tested semiconductor’s chip as the latter is replaced with a new semiconductor’s chip. If external factors cause the immediate replacement of a new chip after the semiconductor’s chip being tested is completed, then the indexing time is rejected. The test time was considered valid if no external factors inclusive of the tester downtime and chip contacting problems cause a high rejection rate of the tested chip.

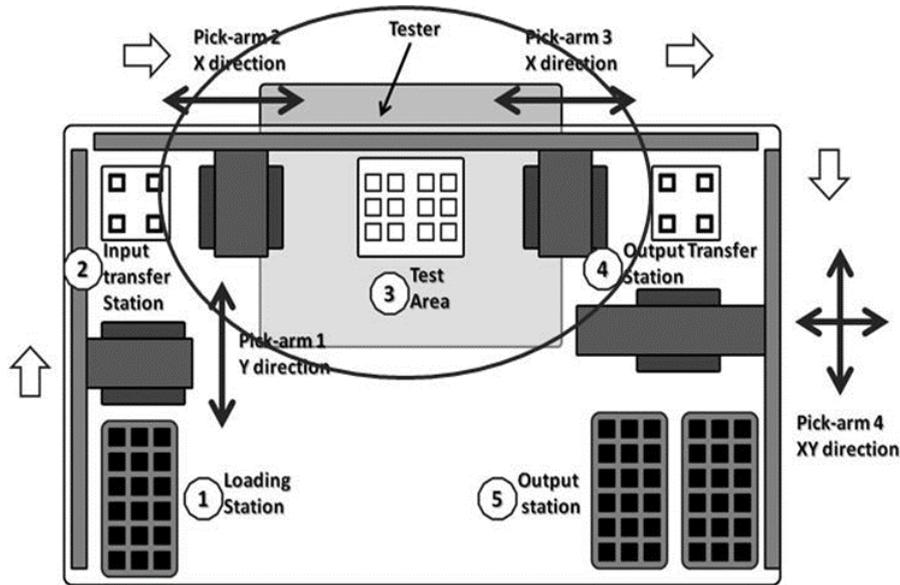


Fig. 4. Focus Area of the Research

4.3 Research hypotheses

The data were then analyzed with SPSS21 for Pearson correlation coefficient test, and the following research hypotheses were proposed in this study:

H0 1: No significant inter-correlations exist between velocity and UPH

H0 2: No significant inter-correlations exist between velocity and the profit

H0 3: No significant inter-correlations exist between velocity and the ROI

H0 4: No significant inter-correlations exist between veracity and UPH

H0 5: No significant inter-correlations exist between veracity and the profit

H0 6: No significant inter-correlations exist between veracity and the ROI

The analyses were based on the “strengths of the correlation coefficient values,” as tabulated in Table 3.

Table 3. Strengths of the Correlation Coefficient Values

| Correlation coefficient size (<i>r</i>) | Correlation strength |
|---|----------------------|
| 0.91 to 1.00 or -0.91 to -1.00 | Very strong |
| 0.71 to 0.90 or -0.71 to -0.90 | Strong |
| 0.51 to 0.70 or -0.51 to -0.70 | Average |
| 0.31 to 0.50 or -0.31 to -0.50 | Weak |
| 0.01 to 0.30 or -0.01 to -0.30 | Very weak |
| 0.00 | No correlation |

5 Result and Discussion

5.1 Model validation

To validate the model, Equations 2 and 3 were compared, and the comparison results are tabulated in Table 4.

Table 4. Model Validation

| Equation 2 | | | | | | | | |
|--|--------|--------|--------|--------|-------------------|------|-------------|----------------|
| Single-site indexing time (<i>i</i>) and single-site test time (<i>t</i> ₁) | MSE | MSEAI | Yield% | AIY% | Veracity accuracy | ECR% | Server time | Equation 2 UPH |
| 1.354 | 31.47% | NA | 99.14% | NA | NA | NA | NA | 3,791.82 |
| Equation 3 | | | | | | | | |
| | | | | | | | | Equation 3 UPH |
| 1.354 | NA | 31.47% | NA | 99.14% | 0 | 0 | 0 | 3,791.82 |
| 1.354 | NA | 40.75% | NA | 95.02% | 15% | 5% | 0.2 | 3,395.28 |
| 1.354 | NA | 53.13% | NA | 96.19% | 35% | 5% | 0.3 | 3,689.16 |
| 1.354 | NA | 65.51% | NA | 97.36% | 55% | 5% | 0.45 | 3,823.50 |
| 1.354 | NA | 77.88% | NA | 98.53% | 75% | 5% | 1.15 | 2,515.07 |
| 1.354 | NA | 90.26% | NA | 99.71% | 95% | 5% | 1.6 | 2,149.75 |

As shown in Table 4, when the veracity accuracy (VRA%), environment control error (ERC%), and server time are equal to zero, then MSEAI, AIY%, and Equation 3 UPH are equal to MSE, Yield%, and Equation 2 UPH, respectively. When VRA% increases, then the MSEAI will relatively increase because the multi-site test time was reduced by reduce pin count test corrected by VRA%. However, when the server time increases, then the velocity increases, which relatively slows down the UPH. Table 4 concludes that when VRA%, ERC%, and server time are equal to zero, then, similar with no Industry 4.0 implementation, MSE, Yield%, and Equation 3 UPH are equal to the computation results in Equation 2. Thus, Equation 3 is validated as accurate. Furthermore, Table 5 shows that the inter-correlation analysis results reflect a very strong and positive correlation between UPH and profit ($r = 1.00$) and ROI ($r = 1.00$). All these correlations are significant at $p < 0.01$.

5.2 Discussion of hypothesis

Table 5. Pearson Correlation Coefficient Analysis

| | | Correlations | | | | |
|---------------------|---------------------|-----------------|----------------------------|------------|---------------|-----------------|
| | | <i>Velocity</i> | <i>Veracity accuracy %</i> | <i>UPH</i> | <i>Profit</i> | <i>ROI rate</i> |
| Velocity | Pearson Correlation | 1 | 0.15** | -0.67** | -0.67** | -0.67** |
| | Sig. (2-tailed) | | 0.00 | 0.00 | 0.00 | 0.00 |
| | N | 780 | 780 | 780 | 780 | 780 |
| Veracity accuracy % | Pearson correlation | 0.15** | 1 | 0.50** | 0.50** | 0.50** |
| | Sig. (2-tailed) | 0.000 | | 0.00 | 0.00 | 0.00 |
| | N | 780 | 780 | 780 | 780 | 780 |
| UPH | Pearson correlation | -0.67** | 0.50** | 1 | 1.00** | 1.00** |
| | Sig. (2-tailed) | 0.00 | 0.00 | | 0.00 | 0.00 |
| | N | 780 | 780 | 780 | 780 | 780 |
| Profit | Pearson correlation | -0.67** | 0.50** | 1.00** | 1 | 1.00** |
| | Sig. (2-tailed) | 0.00 | 0.00 | 0.00 | | 0.00 |
| | N | 780 | 780 | 780 | 780 | 780 |
| ROI rate | Pearson correlation | -0.67** | 0.50** | 1.00** | 1.00** | 1 |
| | Sig. (2-tailed) | 0.00 | 0.00 | 0.00 | 0.00 | |
| | N | 780 | 780 | 780 | 780 | 780 |

** . Correlation is significant at the 0.01 level (2-tailed).

The Pearson correlation coefficient analysis results are discussed as follows:

H0 1: No significant inter-correlations exist between velocity and UPH.

The inter-correlation analysis results show an average and negative correlation between velocity and UPH ($r = -0.67$) and significance at $p < 0.01$. Therefore, the directional hypothesis is rejected and the velocity and UPH have average and negative significant inter-correlations.

H0 2: No significant inter-correlations exist between velocity and the profit.

The inter-correlation analysis results show an average and negative correlation between velocity and profit ($r = -0.67$) and significance at $p < 0.01$. Therefore, the directional hypothesis is rejected and the velocity and profit have average and negative significant inter-correlations.

H0 3: No significant inter-correlations exist between velocity and the ROI.

The inter-correlation analysis results show an average and negative correlation between velocity and ROI ($r = -0.67$) and significance at $p < 0.01$. Therefore, the directional hypothesis is rejected and the velocity and ROI have average and negative significant inter-correlations.

H0 4: No significant inter-correlations exist between veracity and UPH.

The inter-correlation analysis results show an average and positive correlation between veracity accuracy percentage and UPH ($r = 0.50$) and significance at $p < 0.01$. Therefore, the directional hypothesis is rejected and the veracity accuracy percentage and UPH have average and positive significant inter-correlations.

H0 5: No significant inter-correlations exist between veracity and profit.

The inter-correlation analysis results show an average and positive correlation between veracity accuracy percentage and profit ($r = 0.50$) and significance at $p < 0.01$. Therefore, the directional hypothesis is rejected and the veracity accuracy percentage and profit have average and positive significant inter-correlations.

H0 6: No significant inter-correlations exist between veracity and the ROI.

The inter-correlation analysis results show an average and positive correlation between veracity accuracy percentage and ROI ($r = 0.50$) and significance at $p < 0.01$. Therefore, the directional hypothesis is rejected and the veracity accuracy percentage and the ROI have average and positive significant inter-correlations.

5.3 Research implications

Firstly, the research results show that the increase in big data velocity affects the overall production output. This finding was supported by the research of [14], where data velocity management involves bandwidth and ingest issues. In other words, when the performance of the server for data management is efficient, then the processing time becomes less. Thus, the selection of a database server type is important to ensure that the implementation of Industry 4.0 will benefit the industry and shorten the return of investment period. As a result, the implementation can be a payoff and Industry 4.0 can prove its title as a revolutionary technology.

Secondly, the research concluded that big data must achieve a high percentage of veracity accuracy to make sure that artificial intelligent and machine learning can produce and predict situations to make accurate decisions. Big data have no value unless they can be effectively utilized, and the proper utilization of big data depends on their IQ [23]. Otherwise, big data might bring negative effects rather than positive effects. Inherent uncertainty in big data veracity has become one of the deciding factors of the successful implementation of Industry 4.0. Therefore, the development of new big data management technology and selection of its application are important in revolutionizing the Industry 4.0 revolution.

5.4 Future works

This research has developed the TI.4ROIR as one of the initial guidelines for future research studies on Industry 4.0. However, the current theory is still imperfect because it only focuses on the semiconductor testing industry. Future research can further enhance the theory so that it can be a unified ROI theory being implemented for all industries or to be implemented in the proposed industry 4.0 revolution. This goal will not be achievable unless case studies are conducted in all the potential industries and further data are collected to prove the robustness of the theory. Therefore, future researchers are encouraged to put more effort into this matter to realize better productivity and profitability in all industries in the common interest of the well-being of the public and the economic wealth-creation of the country. Finally, because we are living in the multimedia and information age, the theory can be further converted into a software application so that similar future researches can be conducted in a more convenient manner.

5.5 Conclusion

Velocity affects UPH, which is the good-unit output of the pick-and-place test equipment. When TIDB increases, velocity slows down, causing a negative impact on UPH and subsequently affects depreciation of profit and slows down ROI rate. In this case, the research concluded that, firstly, choosing the right database type is important to guarantee the success of Industry 4.0 implementation. And secondly, big data are inherently dirty and uncertain. Based on the analysis results, the increasing veracity accuracy of big data will provide a positive impact on UPH because when VRA% increases, then UPH will relatively increase and contribute to the improvement of profit and ROI rate. In conclusion, the proposed theory of TI.4ROIR is deemed to contribute to a new understanding of firm profit and the ROI theory, where the technical aspects of the pick-and-place test equipment and variables of Industry 4.0 are incorporated together. Finally, the theory will assist in Industry 4.0 research in the future.

6 Acknowledgement

The author is thankful to God the Almighty for granting him a healthy body and continued presence of mind when confronting and overcoming the numerous challenges in completing this research. The author hopes that this research will contribute to the betterment of human society as part of God's will for mankind.

7 References

- [1] Ashton, T. S. (1948). Some Statistics of the Industrial Revolution in Britain 1. The Manchester School, 16(2), 214-234. <https://doi.org/10.1111/j.1467-9957.1948.tb00575.x>
- [2] Bhadani, A. K., & Jothimani, D. (2016). Big Data: Challenges, Opportunities, and Realities. In *Effective Big Data Management and Opportunities for Implementation* (pp. 1-24). IGI Global. <https://doi.org/10.4018/978-1-5225-0182-4.ch001>
- [3] Barry, A., & D. Slater., (2005). *The Technological Economy*. London: Routledge.
- [4] Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, M. (2014). How virtualization, decentralization and network building change the manufacturing landscape: An Industry 4.0 Perspective. *International Journal of Mechanical, Industrial Science and Engineering*, 8(1), 37-44.
- [5] Brodman, J. G., & Johnson, D. L. (1995). Return on investment (ROI) from software process improvement as measured by US industry. *Software Process: Improvement and Practice*, 1(1), 35-47.
- [6] Crafts, N. F., & Harley, C. K. (1992). Output growth and the British industrial revolution: a restatement of the Crafts-Harley view 1. *The Economic History Review*, 45(4), 703-730. <https://doi.org/10.1111/j.1468-0289.1992.tb01313.x>
- [7] Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013, May). Addressing big data issues in scientific data infrastructure. In *Collaboration Technologies and Systems (CTS), 2013 International Conference on* (pp. 48-55). IEEE. <https://doi.org/10.1109/cts.2013.6567203>

- [8] Drath, R., & Horch, A. (2014). Industrie 4.0: Hit or hype?[industry forum]. IEEE industrial electronics magazine, 8(2), 56-58. <https://doi.org/10.1109/mie.2014.2312079>
- [9] Evans, A.C. (1999) “Applications of Semiconductor Test Economics, and Multisite Testing to Lower Cost of Test,” Proceedings of the International Test Conference, pp. 113-123. <https://doi.org/10.1109/test.1999.805620>
- [10] Friedlob, G. T., & Plewa Jr, F. J. (1996). Understanding return on investment. John Wiley & Sons.
- [11] Gilchrist, A. (2016). Industry 4.0: the industrial internet of things. Apress.
- [12] Howell, B., & Grimes, A. (2010). Productivity Questions for Public Sector Fast Fibre Network Financiers. Communications & Strategies, 1(78), 127-146.
- [13] Kepner, J., Gadepally, V., Michaleas, P., Schear, N., Varia, M., Yerukhimovich, A., & Cunningham, R. K. (2014). Computing on masked data: a high performance method for improving big data veracity. In High Performance Extreme Computing Conference (HPEC), 2014 IEEE (pp. 1-6). IEEE. <https://doi.org/10.1109/hpec.2014.7040946>
- [14] Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013, January). Big data: Issues and challenges moving forward. In System sciences (HICSS), 2013 46th Hawaii international conference on (pp. 995-1004). IEEE. <https://doi.org/10.1109/hicss.2013.645>
- [15] Kelly, J. (2008) “Multi-site Efficiency and Throughput,” Technical Paper for Verigy.
- [16] Khoo, V. C. (2014). A case study on the effectiveness of Multi-sites test handler to improve of Production Output. IOSR Journal of Engineering, 4(4), 47-59. <https://doi.org/10.9790/3021-04454759>
- [17] Khoo, V. C. (2015). A Cost Of Test Case Study For Wafer-Ring Multi-Sites Test Handler In Semiconductor’s Industry Through Theory Of The Firm. Jurnal Teknologi, 73(1), 101-109. <https://doi.org/10.11113/jt.v73.2855>
- [18] Lee, J., Kao, H. A., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment. Procedia Cirp, 16, 3-8. <https://doi.org/10.1016/j.procir.2014.02.001>
- [19] Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18-23. <https://doi.org/10.1016/j.mfglet.2014.12.001>
- [20] Lasi, H., Fettke, P., Kemper, H. G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. Business & Information Systems Engineering, 6(4), 239-242. <https://doi.org/10.1007/s12599-014-0334-4>
- [21] Li, Y., & Manoharan, S. (2013). A performance comparison of SQL and NoSQL databases. In Communications, computers and signal processing (PACRIM), 2013 IEEE pacific rim conference on (pp. 15-19). IEEE. <https://doi.org/10.1109/pacrim.2013.6625441>
- [22] Lukoianova, T., & Rubin, V. L. (2014). Veracity Roadmap: Is Big Data Objective, Truthful and Credible?. Advances in Classification Research Online, 24(1), 4-15. <https://doi.org/10.7152/acro.v24i1.14671>
- [23] Slaughter, S. A., Harter, D. E., & Krishnan, M. S. (1998). Evaluating the cost of software quality. Communications of the ACM, 41(8), 67-73. <https://doi.org/10.1145/280324.280335>
- [24] Smith, A. (1776) The Wealth of Nations. New York: Xlodern Library.
- [25] Thamhain, H.J. (2005) Management of Technology. New Jersey: John Wiley & Sons, Inc.
- [26] Volkerink E.H, Khoche, A., Rivoir, J. and Hilliges, K..D.(2002) Test Economics for Multi-site Test with Modern Cost Reduction Techniques. Proceedings of the 20th IEEE VLSI Test Symposium. <https://doi.org/10.1109/vts.2002.1011173>
- [27] Vengedassalam, D., & Madhavan, K., (2013). Principles of economics. Malaysia: Oxford Fajar.

8 Author

Khoo Voon Ching is a Doctor of Engineering student in University Technology of Malaysia. Khoo Voon Ching holds two master's degrees, namely, the M.Sc by research from Asia e University and Master of Business Administration from Akamai University, USA. Voon Ching obtained his professional qualifications as an incorporated engineer from the Engineering Council, UK. He also studied in the University Technology of Malaysia to obtain his diploma in Mechanical Engineering and in Institute First Robotics Industrial Science to acquire an advanced diploma in Robotics and Automation Engineering. He obtained his Doctor of Philosophy in Business Administration from Bulacan State University.

Article submitted 2019-06-18. Resubmitted 2019-07-24. Final acceptance 2019-07-25. Final version published as submitted by the authors.